IDAGs: a Perfect Map for Any Distribution

R.R. Bouckaert

UU-CS-1994-22
June 1994

Utrecht University

Department of Computer Science

Padualaan 14, P.O. Box 80.089,
3508 TB Utrecht, The Netherlands,
Tel. : ... + 31 - 30 - 531454

IDAGs:
a Perfect Map for Any Distribution

R.R. Bouckaert

Technical Report UU-CS-1994-22
June 1994

Department of Computer Science
Utrecht University
P.O.Box 80.089
3508 TB Utrecht
The Netherlands

ISSN: 0924–3275

IDAGs: a Perfect Map for Any Distribution

Remco R. Bouckaert

Department of Computer Science, Utrecht University,
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands,
e-mail: remco@cs.ruu.nl

Abstract

Conditional independence relations appear in many different fields of science such as probability theory, theory of belief functions, and possibility theory. Directed acyclic graphs (DAGs) are a powerful means for representing conditional independence relations. However, it is not possible for all independence relations to find a DAG that perfectly represents all conditional independencies in the relation. We present the IDAG formalism that provides for an exact representation of any independency relation. The basic idea is to enhance the formalism of DAGs with a special kind of arc for modeling induced independencies.

1 Introduction

In the last twenty years, conditional independence has been studied in many different areas of computer science and mathematics. Applications of conditional independence can be found in probabilistic inference, theory of relational databases, theory of belief functions, theory of ordinal conditional functions, and possibility theory, [13].

In this paper, we are especially interested in the use of conditional independence in probabilistic inference: reasoning with a joint probability distribution over a set of variables. In probabilistic inference, the volume of information would be unmanageable if no assumptions about independence could be made; the independence assumptions justify an efficient division of the joint probability distribution. The notion of independence is closely related to the notion of relevance and relevance exhibits a monotonic and non-monotonic behavior. When information is gained, previously relevant information may become irrelevant, and on the other hand, previously irrelevant information may become relevant. Conditional independence has a firmly mathematical basis, [3, 12].

The formalism of directed acyclic graphs (DAG) is a powerful means for representing conditional independencies [6, 8, 10, 11]. Furthermore, efficient algorithms are known for probabilistic inference [7, 9, 10] that exploit the represented conditional independencies in DAGs. However, the formalism of DAGs is not expressive enough to perfectly represent all independencies of any distribution [10]. For independence relations that cannot be exactly represented, some irrelevancies are not reflected in the DAG causing an unnecessary computational burden in probabilistic inference. Extended formalisms have been proposed to capture all independencies of an independence relation in a collection of graphs [5, 6]; however, inference with a collection of graphs is still computational expensive.

In this paper, a new formalism for representing independencies called IDAGs is presented that provides for exact modelling of all independencies of any positive distribution, or to be more precisely, of any independence relation that satisfies the graphoid axioms. The basic idea is to enhance the formalism of DAGs with a new type of arc that allows for representing induced independencies. These arcs closely resemble the justifications used in truth maintenance systems, [4]: the arcs specify a monotonic part as well as a non-monotonic part. In Section 2, we present some preliminaries concerning conditional independence and DAGs. In Section 3, we introduce the IDAG formalism. We end with conclusions and some suggestions for further research in Section 4.

2 Conditional independence

Let U be a non-empty finite set of statistical variables. In the sequel, we will use capital letters to denote sets of variables and lower-case letters to denote single variables; to prevent abundant usage of symbols, we write x to denote $\{x\}$, XY to denote $X \cup Y$, xy to denote $\{x, y\}$, and xY to denote $\{x\} \cup Y$. Further, we let union be of higher priority than exclusion; we write $WX\backslash yz$ to denote $(W \cup X)\backslash(\{y, z\})$. Let X, Y, Z be disjoint subsets of U where $X \neq \emptyset$, $Y \neq \emptyset$. We write $I(X, Z, Y)$ to denote that X and Y are *conditionally independent* given Z; $I(X, Z, Y)$ is called an *independency statement* over U. An *independence model* is a set of independency statements. A *graphoid independence model*, is an independency model that is closed under the the following rules called *graphoid axioms* or *independency axioms*,

symmetry	$I(X, Z, Y)$	$\Leftrightarrow I(Y, Z, X)$
decomposition	$I(X, Z, WY)$	$\Rightarrow I(X, Z, W)$
weak union	$I(X, Z, WY)$	$\Rightarrow I(X, ZW, Y)$
contraction	$I(X, ZW, Y) \wedge I(X, Z, W)$	$\Rightarrow I(X, Z, WY)$
intersection	$I(X, ZW, Y) \wedge I(X, ZY, W)$	$\Rightarrow I(X, Z, WY)$

The graphoid axioms can be considered rules of inference for graphoid independency models. For example, if $I(X, Z, WY)$ is in a model then, by the decomposition axiom, $I(X, Z, Y)$ is also in the model. In this paper, we consider graphoid independency models only: every independency model mentioned is assumed to be graphoid unless stated otherwise.

Graphs offer a powerful means for representing independency models. We review some basic definitions from graph theory. A ***directed graph*** G is a pair (U, A) where U is a set of nodes and A is a set of ordered tuples (x, y) called arcs where $x, y \in U$. In the sequel, an arc (x, y) is denoted by $x \rightarrow y$ or by $y \leftarrow x$, alternatively. A ***path*** in a directed graph G is an ordered set of arcs $x_1 \rightarrow y_1, x_2 \rightarrow y_2, \ldots, x_k \rightarrow y_k$ such that $y_i = x_{i+1}$ for all $i = 1, \ldots, k-1$; the path is called a ***cycle*** if $x_1 = y_k$. A ***directed acyclic graph*** (DAG) is a directed graph that does not contain cycles. A ***trail*** in a DAG is a path that does not consider the direction of the arcs. A ***head-to-head node*** on a trail is a node x such that a triple of consecutive nodes y, x, z $(y \neq z)$ are on the trail and arcs $y \rightarrow x$, $x \leftarrow z$ in the DAG.

DAGs can be used for the representation of independency statements. With every variable $u \in U$ a node in the DAG is associated. Independency statements can be read from the structure of the DAG with the notions of blocked and separation [6, 10].

Definition 2.1 *Let G be a DAG. A trail in G between two nodes x and y is* blocked *by a set of nodes Z if at least one of the following two conditions holds:*
- *the trail contains a node x such that $x \in Z$ and x is not a head-to-head node on the trail;*
- *the trail contains a node x such that x is a head-to-head node on the trail, and x and every descendant of x is not in Z.*

Definition 2.2 *Let G be a DAG. Let X, Y, and Z be sets of nodes of G. We say that X is* d-separated *from Y given Z in G, written $<X, Z, Y>_G$, if every trail between any node $x \in X$ and any node $y \in Y$ is blocked by Z in G; $<X, Z, Y>_G$ is called a* d-separation statement.

In the sequel, we will write $<X, Z, Y>$ instead of $<X, Z, Y>_G$ as long as no ambiguity occurs. For d-separation statements, the graphoid axioms apply, [10]; for example, if $<X, Z, WY>$ holds, then $<X, Z, W>$ also holds by the decomposition axiom.

Definition 2.3 *Let G be a DAG and let M be an independency model. G is an* independency map *or* I-map *of M if $<X, Z, Y>$ in G implies $I(X, Z, Y) \in M$. An independency map of M is* minimal *if no proper subgraph of G is an independency map of M. G is a* dependency map *or* D-map *of M if $I(X, Z, Y) \in M$ implies $<X, Z, Y>$ in G. G is a* perfect map *or* P-map *of M if it is both an independency map and a dependency map of M.*

Not every independency model has a perfect map. Consider for example, the well known coins and bell experiment, [10]; two coins are flipped and a bell rings if the outcomes of both coins is the same. In this experiment we have three variables a, b, and c. These variables are pairwise independent, so the independency model M comprises the statements $I(a, \emptyset, b)$, $I(a, \emptyset, c)$, $I(b, \emptyset, c)$, and symmetric statements. Yet, given the third variable any

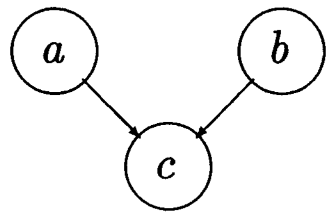

Figure 1: A DAG representing the coins and bell experiment.

pair of variables is dependent. This property cannot be perfectly represented by a DAG. For a statement $<x, \emptyset, y>$ to hold in a DAG, there cannot be an arc between x and y. So, a DAG representing the coins and bell experiment should not have any arcs; yet, such a DAG would represent $I(a, b, c)$, a statement that is not in the model.

It is better to represent fewer independencies than in a model by a DAG than to represent some independencies that are not in the model; representing more independencies would imply adding more irrelevancies and may cause erroneous conclusions in for example probabilistic inference. So, the best we can do is to represent a model by a minimal I-map. For the model of the coins and bell experiment, Figure 1 depicts an example of a minimal I-map. A minimal I-map can be constructed from an independency model M using the notion of causal input list [11].

Definition 2.4 *Let M be a graphoid independency model over U. Let θ be a total ordering on U. A* causal input list *L_θ over M is a set of independency statements such that for each $x \in U$ the set contains a statement $I(x, \pi_x, U_x \backslash \pi_x)$ where $U_x = \{y | \theta(y) < \theta(x)\}$ and π_x is the smallest subset of U_x for which $I(x, \pi_x, U_x \backslash \pi_x) \in M$. π_x is called the* parent set *of x.*

From a causal input list L_θ over an independency model M, a DAG G_θ is constructed by taking for the parent sets of the nodes the corresponding parent sets in L_θ. The thus constructed DAG G_θ is a minimal I-map of M, [11]. The model M_θ represented by G_θ by d-separation is equal to the closure of L_θ under the graphoid axioms (follows from [11]).

An application of the theory of independency models is employed in probabilistic networks. A *probabilistic network* is a pair (G, Γ) where G is a DAG and Γ is a set of *conditional probability tables* $P(x|\pi_x)$ for every variable $x \in U$. A conditional probability table enumerates the probabilities of a given variable given the values of its parents in the DAG. The conditional probability tables taken together define a probability distribution that respects the independency model represented by the DAG, that is, if X and Y are d-separated by Z in G, then the independency statements $I(X, Z, Y)$ holds in the distribution represented by the network, [6, 10].

3 IDAGs

In this section, we present a new formalism, called the *IDAG formalism*, for representing independency models. We begin by introducing a new type of arc called independency inducing arc.

Definition 3.1 *Let $G = (U, A)$ be a DAG. An* independency inducing arc *(IIA) on G is an ordered triple γ, α, β, written $(\alpha \mid \beta \rightarrow \gamma)$, where $\gamma = (x, y)$ is an unordered pair of nodes and $\alpha, \beta \subseteq U \backslash xy$ are two disjoint sets of nodes of G; α is called the* monotonic support *of γ and β is called the* non-monotonic support *of γ.*

A DAG can be extended with these arcs, resulting in an IIA-directed acyclic graph (IDAG).

Definition 3.2 *Let $G' = (U, A)$ be a DAG. An* IIA-directed acyclic graph *(IDAG) is an ordered tuple $G = (U, A, Q)$, where Q is a set of independency inducing arcs on G'; G' is called the* embedded DAG *of G.*

The independency inducing arcs are modelled after the justifications in truth maintenance systems [4]. An independency inducing arc $(\alpha \mid \beta \rightarrow (x, y))$ is taken to represent that the nodes x and y are independent given a set of nodes Z if all nodes of the monotonic support α are in Z and no node of the non-monotonic support β is in Z; if at least one of these conditions does not hold, then the arc does not indicate a change to the independencies represented by the embedded DAG. Note that independency inducing arcs may exists for nodes that are not adjacent in the embedded DAG. For example, a DAG $a \rightarrow b \leftarrow c$ may contain an independency inducing arc $(b \mid \emptyset \rightarrow (a, c))$. As for DAGs, this meaning is formulated in a criterion for reading independency statements from an IDAG.

Definition 3.3 *Let $G = (U, A, Q)$ be an IDAG. A trail between two nodes x and y in the embedded DAG G' is* I-blocked *by a set of nodes Z in G if at least one of the following conditions holds:*
- *the trail is blocked by Z in G';*
- *there is an independency inducing arc $(\alpha \mid \beta \rightarrow (x, y))$ in Q such that $\alpha \subseteq Z$ and $\beta \cap Z = \emptyset$.*

Definition 3.4 *Let G be an IDAG. Let X, Y, and Z be sets of nodes of G. We say that X is* I-separated *from Y given Z in G, written $<X, Z, Y>_G$ if every trail between any node $x \in X$ and any node $y \in Y$ in the embedded DAG of G is I-blocked by $XYZ \backslash xy$. $<X, Z, Y>_G$ is called an* I-separation statement.

Note that we use the same notation for I-separation as for d-separation. The context will make clear which kind of separation is meant. I-separation can be regarded an extension of d-separation. Let G' be a DAG and G be an IDAG that has G' as its embedded DAG.

Suppose that G does not contain any independency inducing arc. If $<X, Z, Y>_{G'}$, then also for all $x \in X$ and $y \in Y$, $<x, XYZ\backslash xy, y>_{G'}$, because d-separation obeys weak union and symmetry, [10]. But $<x, XYZ\backslash xy, y>_{G'}$ for all $x \in X$ and $y \in Y$, implies that X and Y are I-separated by Z in G. For an IDAG that does not contain any independency inducing arc, the set of I-separation statements is equal to the set of d-separation statements in the embedded DAG.

The notions of I-map, D-map and P-map for IDAGS are defined similar to the way it has been done for DAGs.

Definition 3.5 *Let G be a IDAG and let M be an independency model. G is an* independency map *or* I-map *of M if $<X, Z, Y>$ in G implies $I(X, Z, Y) \in M$. An independency map of M is* minimal *if no proper subgraph of G is an independency map of M. G is a* dependency map *or* D-map *of M if $I(X, Z, Y) \in M$ implies $<X, Z, Y>$ in G. G is a* perfect map *or* P-map *of M if it is both an independency map and a dependency map of M.*

We give three examples of independency models that cannot be perfectly represented by a DAG an represent them by an IDAG. Consider once more the independency model of the coins and bell example in Section 2. The IDAG shown in Figure 2 is a perfect map of this model. The IDAG has embedded the DAG of Figure 1 and comprises two independency inducing arcs, namely $(\emptyset \mid a \rightarrow (b, c))$ and $(\emptyset \mid b \rightarrow (a, c))$; in the figure an independency inducing arc is depicted as an edge labeled with a rectangle specifying the monotonic and non-monotonic support separated by a '|'. Note that in the IDAG not only $I(a, \emptyset, b)$ but also $I(a, \emptyset, c)$ and $I(b, \emptyset, c)$ (and their symmetric statements) are represented. The dependencies in the model, for example $\neg I(a, b, c)$ and $\neg I(a, c, b)$, are also represented by the IDAG: $<a, b, c>$ and $<a, c, b>$ are no valid I-separation statements.

Consider the venereal disease example described in [10]; two males and two females have mutual heterosexual contact. Let the variables a, b, c, and d represent whether the persons a, b, c, and d have a venereal disease, where a and c are male and b and d are female. The independency model of this example is defined by the statements $I(b, ac, d)$ and $I(a, bd, c)$. This model cannot be perfectly represented by a DAG, [10]. Figure 3 however, shows an IDAG that is a perfect map of the model.

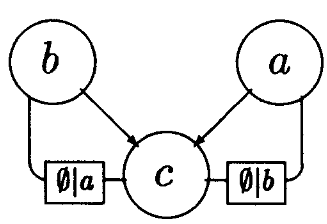

Figure 2: An IDAG that is a P-map for the coins and bell experiment.

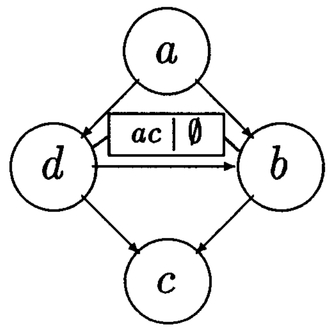

Figure 3: An IDAG that is a perfect map for the venereal disease example.

Another independency model that cannot be perfectly represented by a DAG is the model that is obtained as follows. Let the DAG in Figure 4 be a perfect map of an independency model of a given probability distribution. Then, for the distribution marginalized over c no DAG exists that is a perfect map for the independency model over the four remaining variables. The IDAG depicted in the figure, however, does represent this independency model perfectly. In fact, for every graphoid independency model an IDAG exists that is a perfect map of this model. To proof this, we need the following lemma.

Lemma 3.1 *Let M be a graphoid independency model over U. Let $X, Y, Z \subseteq U$. If for all $x \in X$ and $y \in Y$ we have $I(x, XYZ\backslash xy, y) \in M$, then $I(X, Z, Y) \in M$.*

Proof: Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_m\}$, $n \geq 1$, $m \geq 1$. Now suppose that for all $x_i \in X$, $y_j \in Y$, $i = 1, \ldots, n, j = 1, \ldots, m$ the statement $I(x_i, XYZ\backslash x_i y_j, y_j)$ is valid. Then, for all $x_i \in X$ we have

$I(x_i, XYZ\backslash x_i y_1, y_1)$
$\Rightarrow$ { Intersection with $I(x, XYZ\backslash xy_2, y_2)$ }
$I(x_i, XYZ\backslash x_i y_1 y_2, y_1 y_2)$
$\Rightarrow$ { Intersection with $I(x, XYZ\backslash xy_3, y_3)$ }
$I(x_i, XYZ\backslash x_i y_1 y_2 y_3, y_1 y_2 y_3)$

...

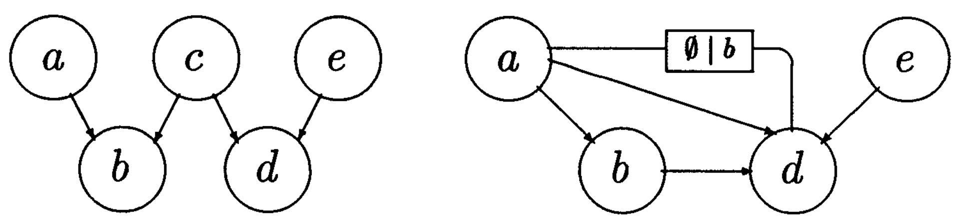

Figure 4: A DAG that is a perfect map and an IDAG that is a P-map for $\{a, b, d, e\}$.

$I(x_i, XYZ\backslash x_i y_1 y_2 \dots y_{m-1}, y_1 y_2 \dots y_{m-1})$
$\Rightarrow$ { Intersection with $I(x, XYZ\backslash xy_m, y_m)$ }
$I(x_i, XZ\backslash x_i, Y)$
$\Rightarrow$ { Symmetry }
$I(Y, XZ\backslash x_i, x_i)$

Applying a similar argument to $x_1, \dots, x_n$, we find $I(X, Z, Y)$. Since M is a graphoid independency model, it satisfies the graphoid axioms. Therefore, we conclude that $I(X, Z, Y) \in M$. □

Theorem 3.1 *Let M be a graphoid independency model. Then, there exists an IDAG G such that G is a perfect map of M.*

Proof: Consider the following construction of an IDAG for an independency model M. First construct any fully connected DAG. Then, add for every independency statement $I(x, Z, y) \in M$, an independency inducing arc $(Z \mid (U\backslash Z)\backslash xy \rightarrow (x, y))$. We show that the IDAG G obtained by this construction is a perfect map of M, that is, $I(X, Z, Y) \in M \Leftrightarrow <X, Z, Y>_G$.

First, assume that $I(X, Z, Y) \in M$ for some $X, Y, Z \subseteq U$. From the weak union and symmetry axioms, we have that $I(X, Z, Y)$ implies that for all $x \in X$ and $y \in Y$ the statement $I(x, XYZ\backslash xy, y)$ holds. By method of construction, for every statement $I(x, XYZ\backslash xy, y)$ the IDAG G comprises an independency inducing arc $(XYZ\backslash xy \mid U\backslash XYZ \rightarrow (x, y))$. Therefore, by definition of I-separation, we have that $<X, Z, Y>_G$.

Now assume that $<X, Z, Y>_G$. By method of construction and from the definition of I-separation, we have that for all $x \in X$ and $y \in Y$ there is an independency inducing arc in G of the form $(XYZ\backslash xy \mid U\backslash XYZ \rightarrow (x, y))$ in G. But then, for all $x \in X$ and $y \in Y$ the statement $I(x, XYZ\backslash xy, y)$ holds. Using Lemma 3.1, we conclude that $I(X, Z, Y) \in M$. □

Consider once more, the IDAG construction mentioned in the above proof. This construction would take $\binom{n}{2}2^{n-2}$ consultations of a given independency model M and would result in an IDAG with a huge number of arcs. A computationally less complex way to construct an IDAG which is a perfect map of the given independency model M is to first generate a DAG G_θ from a causal input list L_θ of M and then add independency inducing arcs $(Z \mid (U\backslash Z)\backslash xy \rightarrow (x, y))$ for every statement $I(x, Z, y)$ that is not yet represented by the IDAG. In other words, an I-map of an independency model M can safely be extended with independency inducing arcs of the proper form, as shown in the following lemma.

Lemma 3.2 *Let M be a graphoid independency model over U. Let G' be a DAG that is an I-map of M. Now, let G be an IDAG such that G' is embedded in G and G comprises independency inducing arcs $(\alpha \mid \beta \rightarrow (x,y))$ only if for all $\alpha \subseteq S \subseteq U \backslash xy\beta$ we have that $I(x,S,y) \in M$. Then, G is an I-map of M*

Proof: Suppose that $<X,Z,Y>$ in G. From the definition of I-separation it follows that for all $x \in X$ and $y \in Y$, $<x, XYZ\backslash xy, y>$ in G; this statement follows either from d-separation of x and y by $XYZ\backslash xy$ in the embedded DAG G', or from the presence of an independency inducing arc $(\alpha \mid \beta \rightarrow (x,y))$ with $\alpha \subseteq S \subseteq U\backslash xy\beta$. In the former case, we have $I(x, XYZ\backslash xy, y)$ is in M because the underlying DAG is an I-map of M; in the latter case, we have this independency statement from the conditions for G stated in the lemma. By Lemma 3.1, we have that $I(X,Z,Y) \in M$. □

In the construction of an IDAG for a given independency model M starting with DAG G' that is a minimal I-map, not all possible independency inducing arcs that could be generated from M need to be inserted in G'; if two nodes x and y are already d-separated by Z in G', then the independency inducing arc $(Z \mid U\backslash xyZ \rightarrow (x,y))$ does not add any new independency statement not yet represented by G'. The fact that the embedded DAG is a minimal I-map can be used to exclude more independency inducing arcs to be considered. Certain separation statements cannot hold in a DAG that is a minimal I-map; if they would hold, it implies an independency statement with which it can be shown that the DAG is not a minimal I-map. For example, let $U = \{a, b\}$ and let G be a DAG with one arc $a \rightarrow b$ being a minimal I-map of an unknown independency model M. Then $I(a, \emptyset, b)$ cannot be in M: if it were in M, then the arc $a \rightarrow b$ could be omitted and G would not be a minimal I-map of M. In [2] a graphical criterion has been presented to read some of these statements from a DAG that is a minimal I-map.

The following rules can be used to reduce the number of arcs of an IDAG that has embedded an I-map of the independency model represented; in presenting the rules we use D_x to denote the set of descendants of node x and π_x to denote the set of parents of x in the embedded DAG of the IDAG at hand.

- *Intersection:*
$(\alpha \cup \{z\} \mid \beta \rightarrow (x,y))$ and $(\alpha \mid \beta \cup \{z\} \rightarrow (x,y))$ can be replaced by $(\alpha \mid \beta \rightarrow (x,y))$.
- *Arc removal:*
$y \rightarrow x$ can be removed if there exist an IIA $(\alpha \mid \beta \rightarrow (x,y))$ in G with $\alpha \cap D_x = \emptyset$ and $\beta \cap \pi_x = \emptyset$.
- *IIA removal:*
$(\alpha \mid \beta \rightarrow (x,y))$ can be removed if $\pi_x \subseteq \alpha$ and $D_x \subseteq \beta$.

The idea of these rules is that application results in a new IDAG representing an independency model implied by the original IDAG, yet with fewer arcs. The new independency model need not be exactly the same as the original model. However, all independency statements in the new model follow from the independency statements in the model before application of one of the rules and the graphoid axioms.

The intersection rule is justified by the fact that an I-separation statement that is satisfied through the independency inducing arc $(\alpha \mid \beta \rightarrow (x,y))$ is either satisfied by the independency inducing arc $(\alpha \cup \{z\} \mid \beta \rightarrow (x,y))$ or by $(\alpha \mid \beta \cup \{z\} \rightarrow (x,y))$. Arc removal of the arc $y \rightarrow x$ with the appropriate independency inducing arc is justified by the fact that the independency inducing arc implies that the represented independency model is implied by the independency statements represented in the original IDAG. Independency inducing arc removal is valid because all trails I-blocked by the independency inducing arc $(\alpha \mid \beta \rightarrow (x,y))$ are also blocked in the embedded DAG. Applying the transformation rules, therefore, does not change the model represented by the IDAG if the IDAG is a P-map, as is stated more formally in the following theorem.

Theorem 3.2 *Let M be a graphoid independency model and let G be an IDAG that is a perfect map of M. Now, let G' be an IDAG obtained from G by applying the intersection, arc removal, or independency inducing arc removal rule shown above. Then, G' is a perfect map of M.*

Proof: From Lemma 3.2 we have that to assure that G' is an I-map of M it is sufficient to show that after application of one of the rules, the embedded DAG is an I-map of M and G contains only independency inducing arcs $(\alpha \mid \beta \rightarrow (x,y))$ such that for all $\alpha \subseteq S \subseteq U \backslash xy\beta$ $I(x,S,y)$ is in M. Further, to assure that G' is a D-map of M we have to show that the independencies represented in G are still represented by G'. From these two properties, we then have that G' is a perfect map of M.

The validity of the intersection rule follows from the definition of I-separation. First, we observe that the presence of the independency inducing arc $(\alpha \cup \{z\} \mid \beta \rightarrow x, y)$ implies for each set S such that $\alpha \cup \{z\} \subseteq S \subseteq U \backslash \beta$ that $I(x,S,y) \in M$. Equally, $(\alpha \mid \beta \cup \{z\} \rightarrow (x,y))$ implies for each set S such that $\alpha \subseteq S \subseteq U \backslash (\beta \cup \{z\})$ that $I(x,S,y) \in M$. So, for each set S such that $\alpha \subseteq S \subseteq U \backslash \beta$ we have that $I(x,S,y) \in M$. This property is also represented by the single independency inducing arc $(\alpha \mid \beta \rightarrow (x,y))$. Not that the embedded DAG is not affected by the applying intersection rule. From these observations we have that after application of the intersection rule G' is a P-map of M

The validity of the arc-removal rule follows from the following observation. Let G'' be the DAG embedded in G. Then, G'' is an I-map of M. Let $U_x = U \backslash x D_x$ for all $x \in U$. Let θ be a topological ordering on G'' thus defining the sets U_x for each node x. Then, for all $x \in U$ we have $I(x, \pi_x, U_x \backslash \pi_x) \in M$. From the presence of the independency inducing arc $(\alpha \mid \beta \rightarrow (x,y))$ in G with $\alpha \cap D_x = \emptyset$ and $\beta \cap \pi_x = \emptyset$, we have that $I(x, U_x \backslash y, y) \in M$. Using intersection, we we find that $I(x, \pi_x \backslash y, (U_x \backslash \pi_x) y)$. Using the causal input list construction with ordering θ results in a DAG G_θ that has for node x a parent set that does not contain

y. We know that $G_\emptyset$ is an I-map of M, [11]. Adding arcs, thus obtaining the embedded DAG of G', still results in an I-map. So, G' is an I-map of M by Lemma 3.2. Observe that no dependencies are added by removing the arc $y \rightarrow x$ and that G was supposed to be a P-map. Therefore, G' also has to be a P-map.

The validity of the independency inducing arc-removal rule follows from the following observation. Consider an independency inducing arc $(\alpha \mid \beta \rightarrow (x, y))$ such that $\pi_x \subseteq \alpha$ and $D_x \subseteq \beta$. By definition of independency inducing arcs, y cannot be in α nor in β and thus not in π_x nor in D_x. Now observe that every trail between x and y in the embedded DAG of G is blocked by any set S such that $\pi_x \subseteq S$ and $S \cap D_x = \emptyset$ since y is not a descendant of x. In any independency statement $I(x, S, y)$ represented by the independency inducing arc the set S contains π_x and $S \cap D_x = \emptyset$. Such statements however are already represented in the embedded DAG. So, no independencies are removed by removing the independency inducing arc. In addition, the embedded DAG is not affected. Hence after independency inducing arc-removal, G' is a P-map of M. □

Note that if the rules are applied to an IDAG that is an I-map of a model M but not a perfect map, the model may change due to application of the arc removal rule. For example, an IDAG over $\{a, b, c\}$ with two arcs $a \rightarrow b$ and $b \rightarrow c$ and an independency inducing arc $(\emptyset \mid \emptyset \rightarrow (a, b))$ may be transformed using arc removal; the arc $a \rightarrow b$ can be removed. Then, $<a, \emptyset, c>$ holds in the new IDAG while it did not in the original IDAG. Application of the two other rules does not influence the represented model.

A major drawback of the IDAG formalism is that an IDAG over a set of n variables may contain an exponential in the number of arcs, that is, exponential in n. Therefore, graphical representation may become infeasible for large networks.

4 Conclusions and Further Research

We have presented the IDAG formalism for perfectly representing any graphoid independency model. The formalism basically enhances the DAG formalism with the notion of independency inducing arc. Similar to the d-separation criterion for DAGs, we have introduced the I-separation criterion for reading independency statements from an IDAG. IDAGs can be used for probabilistic inference by enhancing them with probability tables in the same way DAGs are enhanced with probability tables to obtain a probabilistic network. Efficient algorithms for probabilistic inference based on lazy evaluation of the network may be developed that exploit the independencies induced by independency inducing arcs. It is well-known that algorithms for probabilistic inference with a belief network perform better from a computational point of view when more independencies are represented. As the IDAG formalism allows for more independencies to be represented, especially in large IDAGs the enhanced algorithms for inference may yield a computational saving over algorithms currently known for inference in probabilistic networks.

IDAGs can also be helpful for constructing a DAG from a given independency model. For the construction of a DAG from an independency model, in [1] an algorithm has been presented that optimizes a total ordering θ of a set of variables by applying simple operations to θ. Changing such an ordering implies a change in the causal input list generated from the ordering, and thus a change in the associated DAG G_θ and the represented independency model M_θ. The strategy is to apply an operation to θ only if the represented independency model after the operation contains the one before the operation. This strategy has been shown to yield a DAG with a minimal number of arcs in a local minimum. In order to find the global minimum, we believe the IDAG formalism can be of help, since after an operation on the total¡ ordering θ, independency inducing arcs can be added in order to let the new independency model contain the old one.

Conditional independence relations are used to reflect irrelevancy in knowledge based systems. The notion of irrelevancy that is reflected by conditional independence is global in the sense that it applies to all values of variables. However, variables may be independent for certain combinations of values, yet dependent for other combinations of values. It may be worth while to look for formalisms and efficient inference and explanation algorithms that explore these independencies; introduction of a special kind of arc similar to the independency inducing arc employed by the IDAG formalism may be used to reflect such independencies.

Acknowledgement

I thank Linda van der Gaag for her many helpfull remarks that improved the presentation of this paper considerably.

References

[1] R.R. Bouckaert. Optimizing causal orderings for generating DAGs from data. In *Proceedings Uncertainty in Artificial Intelligence 8*, pages 9–16, 1992.

[2] R.R. Bouckaert. Conditional dependencies in probabilistic networks. In *Proceedings 4th International Workshop on AI and Statistics*, 1993.

[3] A.P. Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society B*, 41(1):1–31, 1979.

[4] J. Doyle. A truth maintenance system. *Artificial Intelligence*, 12:231–272, 1979.

[5] D. Geiger. Towards the formalization of informational dependencies. Technical Report CSD-880053, UCLA, Cognitive Systems Laboratory, Computer Science Department, 1988.

[6] D. Geiger. *Graphoids: a qualitative framework for probabilistic inference.* PhD thesis, UCLA, Cognitive Systems Laboratory, Computer Science Department, 1990.

[7] M. Henrion. An introduction to algorithms for inference in belief nets. In *Proceedings Uncertainty in Artificial Intelligence 6*, pages 129–138, 1990.

[8] S.L. Lauritzen, A.P. Dawid, A.P. Larsen, and H.G. Leimer. Independence properties of directed Markov fields. *Networks*, 20:491–505, 1990.

[9] S.L. Lauritzen and D.J. Spiegelhalter. Local computations with probabilities on graphical structures and their applications to expert systems (with discussion). *Journal of the Royal Statistical Society B*, 50:157–224, 1988.

[10] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufman, inc., San Mateo, CA, 1988.

[11] J. Pearl, D. Geiger, and T. Verma. The logic of influence diagrams. In R.M. Oliver and J.Q. Smith, editors, *Influence Diagrams, Belief Nets and Decision Analysis*, pages 67–87. John Wiley & Sons Ltd., 1990.

[12] M. Studený. Multiinformation and the problem of characterization of conditional-independence relations. *Problems of Control and Information Theory*, 18(1):3–16, 1989.

[13] M. Studený. Formal properties of conditional indpendence in different calculi of AI. In M. Clarke, R. Kruse, and S. Moral, editors, *Proceedings ECSQARU*, volume 747 of *Lecture Notes in Computer Science*, pages 341–348. Springer-Verlag, Berlin, 1993.